Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-328-RC2, 2020

Interactive comment on "On a new assessment method for long-term chemistry-climate simulations in the UTLS based on IAGOS data: application to MOCAGE CCMI-REFC1SD simulation" *by* Yann Cohen et al.

Anonymous Referee #2

Received and published: 31 December 2020

The manuscript presents a new method for comparing the IAGOS observations of O3 and CO taken from in-service aircraft with gridded monthly average fields from chemistry climate models (CCMs) and applies the new method to outputs from the MOCAGE Chemistry Transport Model. The method involves aggregating the in-situ observations, taken along the flight path of the aircraft with a time resolution of less than a minute, to monthly averages at the horizontal and vertical resolution of the model. The method is of considerable interest because collaborative multi-model intercomparison projects frequently request monthly average fields. While requesting monthly average

fields makes the size and complexity of the data request more manageable, comparing monthly average fields with in-situ observations is not straightforward. And the situation is even more complex in the upper troposphere and lower stratosphere due to the strong gradients in chemical species that exist across the tropopause, which is itself highly variable in space and time. By attempting to bridge the gap between the highly time and space resolved IAGOS measurements with the widely available monthly average fields from CCMs the paper represents a nice advance on our understanding of how to assess models.

The paper is relatively well written, though I have noted a few instances where I had difficulty understanding the discussion. In general my comments are minor, detailed below, but I would like to raise one point that I found missing from the discussion of the results. A much better way to compare the models with the IAGOS observations would be to have highly time resolved instantaneous model outputs that are more directly comparable with the in-situ observations. This comparison would not be perfect because of errors in the model that arise from a whole host of different reasons that have been discussed at length in the literature. But, by and large, and particularly so for multi-model intercomparison projects, the large data volumes required to save high frequency output makes that type of comparison difficult and we must work with monthly average data. In addition to the usual list of reasons for model biases, now there is the added factor of having averaged the model fields in time. The authors have done a great job of exploring this effect with the observations by comparing IAGOS-HR with IAGOS-DM, but the discussion of the results does not include much mention of the effect of time averaging. It became a bit confusing when the discussion of the comparison of IAGOS-HR and IAGOS-DM zeroed in on the effect of mis-classification of points in either the UT or LS (see the comment on Page 19, Lines 4 – 6) and ignored the effect of time averaging. In the vicinity of the tropopause, over the course of a month, a particular model grid point will sometimes be in the stratosphere and sometime in the troposphere so the monthly average will reflect both of these influences. At least to my mind, this effect is an important part of the problem comparing monthly-average fields with in-situ observations but I do not find much discussion of this facet of the problem in the manuscript.

Minor comments:

Page 3, Line 28: When discussing the available frequency of model output to compare against the IAGOS data, the manuscript states 'the 3D outputs from the REF-C1SD simulations, which are monthly averages.' The fact that monthly average fields are the most commonly available output is not necessarily part of the REF-C1SD specification, but more of a result of asking for a large amount of data from a number of models participating in a multi-model intercomparison. The text here should be more general and refer to the fact that monthly average fields are the most commonly available type from multi-model intercomparisons.

Page 4, Lines 7 – 10: In discussing other available methods for comparing in situ measurements with model data, the authors state that interpolating neighbouring measurement points onto each gridpoint would be computationally expensive for the IAGOS data because it requires keeping track of a large number of measurement locations each month. But when the methodology is described in Section 3 it would seem to me that the proposed method also requires keeping track of the discrete locations of a large number of observations. I believe I see the idea the authors are trying to convey – that a month of observation data at the original locations must be collected up before interpolating on to the grid point - but I would suggest rewording this point.

Page 7, Line 4: The phrase 'measured a mixing ratio Cobs(X) for an X species' might be better as 'measured a mixing ratio Cobs(X) for species X'

Page 8, Line 3: Where it is written 'increasing linearly with the distance between the measurement point and the (i, j, k) grid point.', I think should be 'decreasing linearly with the distance'.

Page 11, Line 1: I find the phrase 'before deriving monthly means in the two layers'

a bit confusing because I am not familiar with the analysis performed in Cohen et al. (2018) and the earlier discussion of model layers. I assume the two layers are the UT and LS and the analysis is done separately for each region?

Page 11, Line 23: might be missing a word (such?) in 'account, [such] as Western North America and Siberia.'

Page 14, Lines 13-14: 'With respect to the 1:1 line, levels 25 and 26 are characterized by an overestimation of the lower part of the O3 distribution (< 120 ppb) and by an underestimation of the higher part.' This sentence is referring to the pronounced underestimation of ozone between 150 ppb and 250 ppb shown in Figure 5 for P$\sim$255 and P$\sim$285 hPa, and from Figure 3 this looks to be related to an underestimation of ozone at high latitudes. Do the authors have any ideas why ozone in this one region seems to be underestimated to such a large degree? I will note that from the figures in the Appendix the underestimation does seem to be most extreme in the summer months.

Page 18, Lines 17 – 18: In referring to the IAGOS-DM data separated into the UT and LS regions ('In contrast, IAGOS-DM refers to the new product presented in this paper, i.e. the IAGOS data distributed on the model's grid, directly comparable to the simulation.') it might be helpful to the reader to remind them that the data is assigned into either the UT and LS based on the monthly average PV at each model grid point.

Page 19, Lines 4 – 6: Here it is stated ' In other words, by using the monthly mean PV from the simulation, some of the IAGOS measurement points may be attributed to the LS while being in the UT (or in the tropopause layer) and vice-versa.'. Here it sounds as though for the IAGOS-DM dataset individual IAGOS observations are being assigned to either the UT or LS based on their position relative to the monthly average tropopause. I had thought the IAGOS-DM was constructed and then the individual monthly-average values on model grid points were assigned to either the UT or LS (Page 11, Lines 4 – 8)? In this case, a particular monthly average value at a particular model grid point may be affected by a mixture air from both the UT and LS. This is not necessarily a bad thing, as the model monthly averages that are being analyzed will have a similar problem.

Page 19, Figure 7. There are a lot of lines on each panel of Figure 7 (likewise for Figure 8) and making a clear comparison between IAGOS-HR, IAGOS-DM and MOCAGE-M is not easy. I am sympathetic to the need to condense graphics to avoid figures with 20 panels, but I was wondering if the authors would consider adding figures to the appendix that directly compare the three datasets for each of the eight regions? For each region there would be one panel showing the annual cycle from IAGOS-HR, IAGOS-DM and MOCAGE-M for the LS and and a second panel for the UT.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-328, 2020.